

AMMI-BOOTSTRAP NO ESTUDO DA INTERAÇÃO GENÓTIPOS POR AMBIENTES EM SOJA

Maria Clideana Cabral MAIA¹
Natal Antonio VELLO²
Osmir José LAVORANTI³
Carlos Tadeu dos Santos DIAS⁴
Roland VENCOVSKY²
Maurisrael de Moura ROCHA⁵
Glauber Henrique de Sousa NUNES⁶

- RESUMO: Esta pesquisa visou estudar a interação entre genótipos e ambientes (GxE) e a estabilidade fenotípica, com o auxílio da reamostragem *bootstrap* no modelo biométrico AMMI (*Additive Main Effects and Multiplicative Interaction Model*) para avaliação da produtividade de grãos de linhagens experimentais de soja. O material envolveu três populações obtidas de um dialeto parcial 4x4 quatro genitores tolerantes a insetos e quatro cultivares adaptadas. Da geração F₃ a F₁₁ foi praticada seleção entre progênies para produtividade de grãos e tolerância ao fotoperíodo. Foram conduzidos 24 experimentos aleatorizados em blocos com duas repetições subdivididas em conjuntos experimentais com testemunhas comuns, combinando as três populações e oito ambientes. Os ambientes (E1 a E8) combinaram dois locais (Anhembí, Areão), dois anos agrícolas (1999/00, 2000/01) e dois sistemas de manejo (controle intensivo de insetos ou CII e controle ecológico de insetos ou CEI). Concluiu-se que todos os genótipos das três populações foram classificados como estáveis, mas em graus variáveis. Foram formados quatro grupos de genótipos com diferentes níveis de estabilidade, tendo a soma dos dois grupos de estabilidade mais elevada incluído pelo menos 50% das linhagens e duas das três testemunhas; o ambiente E8 (Areão, 2000/01, CEI) foi destacado como de estabilidade máxima e os ambientes E1 (Anhembí, 1999/00, CII) e E3 (Areão, 1999/00, CII) como instáveis. O método AMMI com reamostragens *bootstrap* levou a resultados mais detalhados do que os obtidos apenas com o método AMMI.
- PALAVRAS CHAVE: AMMI-*bootstrap*; interação genótipo x ambiente; *glycine max* (L.).

¹ EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária, CNPAT, Caixa Postal: 3761, CEP: 60511110, Fortaleza, CE, Brasil. E-mail: clideana@terra.com.br.

² Departamento de Genética, Universidade de São Paulo, Campus de Piracicaba – ESALQ/USP, CEP 13418-900, Piracicaba, SP, Brasil. E-mail: naavello@esalq.usp.br // rvencovs@carpa.ciagri.usp.br.

³ EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária, EMBRAPA FLORESTAS, Caixa Postal 319, CEP: 83411-000, Colombo, PR, Brasil. E-mail: osmir@cnpf.embrapa.br.

⁴ Departamento de Ciências Exatas, Universidade de São Paulo, Campus de Piracicaba – ESALQ/USP, CEP 13418-900, Piracicaba, SP, Brasil. E-mail: ctsdias@carpa.ciagri.usp.br.

⁵ EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária, EMBRAPA MEIO-NORTE, CEP: 64006-220, Teresina, PI, Brasil. E-mail: mmrocha@cpamn.embrapa.br.

⁶ Departamento de Ciências Vegetais, Universidade Federal do Semi-Árido – Ufersa, CEP 59625-900, Mossoró, RN, Brasil. E-mail: glauber@ufersa.edu.br.

1 Introdução

O cultivo extensivo da soja, contemplando ambientes diversificados, provoca uma resposta diferencial dos genótipos. Neste sentido, a interação entre genótipos e ambientes GxE representa aspecto relevante no contexto do melhoramento. Assim sendo, Duarte e Vencovsky (1999) afirmam que o entendimento deste fenômeno torna-se imprescindível aos programas de melhoramento que procuram minimizar a inconsistência das características relacionadas à produtividade mediante a variação ambiental, para recomendações mais acertadas.

Adaptações específicas de genótipos a ambientes podem fazer a diferença entre uma boa e uma excelente cultivar (Gauch e Zobel, 1996). Contudo, para que seja possível tirar proveito desses efeitos positivos, de acordo com Brasil (1990) e Duarte e Zimmermann (1995), é preciso dispor de metodologias estatísticas adequadas para se estimar e explorar a interação, permitindo assim recomendações regionalizadas.

Para tanto, o procedimento AMMI (*Additive Main Effects and Multiplicative Interaction Model*) tem se mostrado mais informativo no tratamento da interação GxE do que os métodos usuais existentes. Entretanto, os principais pontos negativos dessa metodologia estão relacionados à dificuldade na interpretação da interação, quando há baixa explicação do primeiro componente principal.

De acordo com Lavoranti (2003) a técnica via reamostragem *bootstrap* no modelo AMMI, em conjunto com representações gráficas de regiões de confiança para a estabilidade e gráficos de dispersões dos escores *bootstrap* em *biplot*, se apresenta como um recurso estatístico que melhora a qualidade para predições das estabilidades fenotípicas, quando comparado ao método tradicional AMMI, com representação gráfica em diagrama de dispersão multivariado (*biplot*).

Esta pesquisa visa estudar a interação entre genótipos e ambientes e a estabilidade fenotípica, com o auxílio da reamostragem *bootstrap* no modelo biométrico AMMI para avaliação de linhagens experimentais de soja, com a finalidade de identificar genótipos que reúnam características de alta produtividade e tolerância a insetos mastigadores e sugadores.

2 Material e métodos

O material genético envolve três tipos de populações obtidas a partir de um dialelo parcial 4x4 desenvolvido com a finalidade de reunir os genes de tolerância/resistência a insetos presentes em quatro genitores (IAC-100, Crockett, Lamar e D72-9601-1), com genes para desempenho agrônomico favorável – principalmente, produtividade de grãos e precocidade – presentes em quatro cultivares adaptadas (BR-6, IAS-5, Davis, Ocepar-4).

A diferenciação principal entre as três populações aconteceu na geração F₂, na qual foram empregados três procedimentos distintos de seleção de plantas individuais, conforme relatado a seguir: PCI (populações com controle de insetos) – originadas da seleção para produtividade de grãos, aplicou-se inseticidas durante o ciclo total para o controle de insetos mastigadores e sugadores; PRIS (populações selecionadas para resistência a insetos sugadores) – plantas que receberam aplicações de inseticidas apenas durante o ciclo vegetativo, de maneira que os insetos sugadores pudessem desenvolver-se, multiplicar-se e atacar as vagens e grãos produzidos pelas plantas F₂; PRIM (populações

selecionadas para resistência a insetos mastigadores) – iniciou-se a aplicação de inseticidas após o florescimento, com a finalidade de controlar insetos sugadores.

A partir da geração F_3 os três tipos de populações foram conduzidos de maneira semelhante, por meio de experimentos realizados em locais diferentes. Em cada geração, promoveu-se seleção principalmente para produtividade de grãos e tolerância ao fotoperíodo.

Os experimentos conduzidos em 1999/00 incluíram 120 linhagens experimentais, na geração F_{10} , sendo 40 linhagens de cada tipo de população (PCI, PRIS e PRIM); os experimentos, conduzidos em 2000/01 envolveram um total de 60 linhagens F_{11} , sendo 20 linhagens selecionadas em cada uma das três populações, com base nos resultados obtidos no ano agrícola anterior.

Os 24 experimentos foram conduzidos no delineamento de blocos ao acaso, com as repetições estratificadas em conjuntos experimentais com testemunhas comuns: IAC-100, Ocepar, IAS-5 e Primavera. A cultivar Primavera foi eliminada nos experimentos do ano agrícola 2000/01. Nos experimentos com controle intensivo de insetos (CII), com base em monitoramento freqüente, promoveu-se a aplicação de inseticidas sempre que se detectaram insetos mastigadores e ou dois percevejos/m² de pano (método do pano de batida). Nos experimentos com controle ecológico de insetos (CEI), promoveu-se aplicações de inseticidas apenas na ocorrência de grande quantidade de danos nas folhas causadas por insetos mastigadores e ou quando a infestação natural atingiu quatro percevejos/m² de pano.

O caráter produtividade de grãos (PG) em gramas/parcela foi avaliado, pelo peso dos grãos maduros colhidos na área útil de cada parcela, após um mínimo de três semanas de secagem em ambiente de laboratório para a padronização da umidade.

Foram considerados fixos os efeitos de populações (linhagens – P1, P2, ... , P20 e testemunhas comuns (IAC-100, OCEPAR-4 e IAS-5), visto que se originaram de diferentes cruzamentos, não representando uma mesma população. O efeito de locais foi considerado fixo, pois os locais foram escolhidos por serem contrastantes em uma série de características, como tipo de solo, fertilidade e topografia; o efeito de manejos também foi considerado fixo por envolver diferenças eleitas *a priori* e poder ter sua condição reproduzida satisfatoriamente; conseqüentemente, as interações entre linhagens e locais (PxL), linhagens e manejos (PxM), manejos e locais (MxL) e linhagens, manejos e locais (PxMxL) foram consideradas fixas. Os efeitos de repetições e anos, e as interações que envolvem estes efeitos, foram tratados de forma aleatória.

O desenvolvimento da metodologia AMMI-*bootstrap* (Lavoranti, 2003) consiste em se executar o método AMMI seguindo a proposta de Gollob, a qual garante pelo menor rigor dos graus de liberdade, um número maior de fatores multiplicativos e, conseqüentemente, capta todo o padrão devido à interação GxE. A partir desse modelo, foram determinadas as estimativas das médias dos genótipos (i) nos ambientes (j) (\hat{Y}_{ij}) livres de interferência de ruídos, e, a partir desse novo conjunto de dados, foram obtidas as matrizes de resíduos ou dos efeitos $(ge)_{ij}$, $\hat{G}E_{G \times E} = (\hat{ge}_{ij})$ representadas por:

$$\hat{G}E_{G \times E} = \begin{vmatrix} \hat{g}e_{11} & \hat{g}e_{12} & \cdots & \hat{g}e_{1e} \\ \hat{g}e_{21} & \hat{g}e_{22} & \cdots & \hat{g}e_{2e} \\ \vdots & \vdots & \cdots & \vdots \\ \hat{g}e_{g1} & \hat{g}e_{g1} & \cdots & \hat{g}e_{ge} \end{vmatrix}$$

com:

$$\hat{g}e_{ij} = \hat{y}_{ij} - \hat{y}_{i.} - \hat{y}_{.j} + \hat{y}_{..}$$

sendo que:

$\hat{y}_{i.}$: é a média dos valores do genótipo i ;

$\hat{y}_{.j}$: é a média dos valores do ambiente j ;

$\hat{y}_{..}$: é a média geral dos valores.

Das matrizes $\hat{G}E_{G \times E}$, foram reamostradas 200 matrizes *bootstrap* para genótipos (GE_g^*), sendo o sorteio com reposição executado nas linhas; e também, foram reamostradas 200 matrizes *bootstrap* para ambientes (GE_e^*), sendo o sorteio com reposição executado nas colunas. Assim, obteve-se:

$$\hat{G}E_{kl}^* = \begin{vmatrix} ge_{11}^* & ge_{12}^* & \cdots & ge_{1e}^* \\ ge_{21}^* & ge_{22}^* & \cdots & ge_{2e}^* \\ \vdots & \vdots & \cdots & \vdots \\ ge_{g1}^* & ge_{g2}^* & \cdots & ge_{ge}^* \end{vmatrix},$$

com: $k = g$ ou $k = e$, para genótipos ou ambientes, respectivamente e $l = (1, 2, \dots, 200)$.

As matrizes *bootstrap* de resíduos foram centradas em relação às médias de linhas, de colunas (W. J. Krzanowski em 2002, citado por Lavoranti, 2003) e média geral, a fim de atender às restrições de identificabilidade da matriz de resíduos.

Assim, de maneira esquemática, foram determinadas as matrizes ($\tilde{G}E_{kl}^*$), denominadas matrizes *bootstrap* de resíduos corrigidos:

$$\tilde{G}E_{kl}^* = \begin{vmatrix} \tilde{g}e_{11}^* & \tilde{g}e_{12}^* & \cdots & \tilde{g}e_{1e}^* \\ \tilde{g}e_{21}^* & \tilde{g}e_{22}^* & \cdots & \tilde{g}e_{2e}^* \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{g}e_{g1}^* & \tilde{g}e_{g2}^* & \cdots & \tilde{g}e_{ge}^* \end{vmatrix},$$

em que:

$$\begin{aligned} \tilde{g}e_{ij}^* &= \hat{g}e_{ij}^* - \frac{1}{e} \sum_{j=1}^e \hat{g}e_{ij}^* - \frac{1}{g} \sum_{i=1}^g \hat{g}e_{ij}^* + \frac{1}{ge} \sum_{i=1}^g \sum_{j=1}^e \hat{g}e_{ij}^* \\ &= \hat{g}e_{ij}^* - \hat{g}e_{i.}^* - \hat{g}e_{.j}^* + \hat{g}e_{..}^* \end{aligned}$$

A partir dessas matrizes foram determinados os efeitos *bootstrap* de genótipos (\tilde{GE}_{kl}^*) e efeitos *bootstrap* de ambientes \tilde{H}_{kl}^{**} , por meio da DVS. Nesse caso, a DVS foi aplicada em cada uma das 600 matrizes *bootstrap* de desvios corrigidos:

$$\begin{aligned}\tilde{GE}_{kl}^* &= U_{kl} S_{kl} V_{kl}' = U_{kl} S_{kl}^{\frac{1}{2}} S_{kl}^{\frac{1}{2}} V_{kl}' \\ \tilde{G}_{kl}^* &= U_{kl} S_{kl}^{\frac{1}{2}} \\ \tilde{H}_{kl}^{**} &= S_{kl}^{\frac{1}{2}} V_{kl}'\end{aligned}$$

em que:

- U_{kl} : é a matriz com vetores *bootstrap* singulares-coluna γ_{ikl} ;
- S_{kl} : é a matriz com valores *bootstrap* singulares $\lambda_{kl} e$;
- V_{kl}' : é a matriz com vetores *bootstrap* singulares-linha α'_{jk} .

No método AMMI, a interação GxE é decomposta pela análise de componentes Principais (*Principal Components Analysis* – PCA) em “n” componentes. O termo IPCA vem de “*Incremental Principal Component Analysis*”, ou seja, IPCA₁ é o efeito de interação do componente principal associado ao primeiro eixo, que retêm a maior parte da variação do efeito da interação, IPCA₂ é o efeito de interação do componente principal, associado ao segundo eixo (Mandel, 1971; Zobel et al. 1988 e Gauch e Zobel, 1996).

Dessa forma, foram obtidos para cada genótipo e ambiente, 200 escores associados a cada matriz *bootstrap*, totalizando 15.400 respostas, por eixo da interação GxE (IPCA₁^{*}, IPCA₂^{*}). Esses resultados foram plotados seguindo o modelo AMMI₂, isto é, em eixos cartesianos IPCA₁^{*} x IPCA₂^{*}. A adoção dessa representação gráfica foi ajustada pelo critério de avaliação “*pos-dictiva*” (Gauch e Zobel, 1988) que visa a obtenção de um alto coeficiente de determinação.

A partir dos escores foram estabelecidas as regiões *bootstrap* de confiança, de 99% para $\bar{\theta}_k^{**} = \left[\overline{IPCA_{1k}^*}, \overline{IPCA_{2k}^*} \right]$, ou seja, para os escores médios de cada genótipo e de cada ambiente, separadamente. As regiões de confiança foram obtidas por Weinberg e colaboradores a partir da estimativa não viesada da matriz de covariâncias $S_{\theta k}^*$, para cada par de coordenadas $\bar{\theta}_{kl}^{**} = \left[IPCA_{1kl}^*, IPCA_{2kl}^* \right]$.

A seguir é apresentado o esquema de Lavoranti (2003):

$$\left(\bar{\theta}_k^{**} - \bar{\theta}_{kl}^{**} \right)' \left(S_{\theta k}^* \right)^{-1} \left(\bar{\theta}_k^{**} - \bar{\theta}_{kl}^{**} \right) \leq \frac{2(B-1)}{B(B-2)} F_{(2, B-2)}^\alpha$$

em que:

$$S_{\theta k}^* = \begin{vmatrix} S_{1k}^{2*} & r_k^* S_{1k}^* S_{2k}^* \\ r_k^* S_{1k}^* S_{2k}^* & S_{2k}^{2*} \end{vmatrix},$$

- B : é o número de amostras *bootstrap* ($B = 200$);
 S_{1k}^{2*} : é a variância *bootstrap* para os escores associados ao $IPCA_1^*$;
 S_{2k}^{2*} : é a variância *bootstrap* para os escores associados ao $IPCA_2^*$;
 r_k^* : é o coeficiente de correlação de Pearson para os escores *bootstrap* obtidos a partir de $(IPCA_1^* \times IPCA_2^*)$;
 $K = (g, e)$ para “g” genótipos e “e” ambientes, $i = 1, 2, \dots, g$; $j = 1, 2, \dots, e$;
 $l = 1, 2, \dots, 200$;
 g : é o número de genótipos;
 e : é o número de ambientes.

Visando medir a precisão das estimativas de estabilidade, foram determinados os coeficientes *bootstrap* de estabilidade, denominados (CBE), pela razão entre o número de pontos $(IPCA_1^* \times IPCA_2^*)$ inclusos na região de predição para $\tilde{\theta}_k^* = [0, 0]$ e o tamanho da amostra *bootstrap* (B):

$$CBE = \# \left[\left(\tilde{\theta}_k^* - \tilde{\theta}_{kl}^* \right)' \left(S_{\theta k}^* \right)^{-1} \left(\tilde{\theta}_k^* - \tilde{\theta}_{kl}^* \right) \leq \frac{2(B-1)}{B(B-2)} F_{(2, B-2)}^\alpha \right] B^{-1} \quad (1)$$

A justificativa teórica para a utilização dessa expressão, foi fundamentada na região de predição para uma única observação, nesse caso, o ponto nulo (0,0). A distribuição T^2 de Hotelling foi baseada numa matriz \mathbf{X} com distribuição normal p-variada, com vetor de médias zero e matriz de covariâncias Σ e \mathbf{Z} , com distribuição de Wishart com ν graus de liberdade e parâmetro Σ . Então $\mathbf{X}'\mathbf{Z}^{-1}\mathbf{X}$ tem uma distribuição $\frac{p}{\nu - p + 1} F_{p, \nu - p + 1}$. No caso das regiões de predição *bootstrap* para a origem, \mathbf{X} é uma única estimativa *bootstrap*. A matriz de covariâncias Σ é calculada por \mathbf{S} , obtida de B amostras *bootstrap*, de forma que $(B-1)\mathbf{S}$ substitui \mathbf{Z} com $\nu = B-1$ e $p = 2$. Substituindo esses valores, obtém-se o numerador da expressão (1).

Os genótipos e ambientes estáveis situam-se próximos à origem, isto é, com escores praticamente nulos para os dois eixos da interação GxE ($IPCA_1$ e $IPCA_2$) (Duarte e Vencovsky, 1999). Assim, para as interpretações das estabilidades via análise gráfica (regiões *bootstrap* de confiança) e numérica (CBE), foram usadas os seguintes critérios, segundo Lavoranti (2003):

- a) para os genótipos e ambientes cujas regiões de confiança *bootstrap* continham o ponto de coordenadas nulas (0,0);

Precisão das estimativas de estabilidade	Classificação da estabilidade
$CBE \geq Q_3$ (terceiro quartil)	Muito Alta
Q_2 (segundo quartil) $\leq CBE < Q_3$	Alta
Q_1 (primeiro quartil) $\leq CBE < Q_2$	Média
$CBE < Q_1$	Baixa

- b) os genótipos e ambientes, cujas regiões de confiança que não continham o ponto de coordenada nula, pertencentes a essa categoria foram considerados instáveis.

A adoção dos quartis como medida de classificação de estabilidade, visou dividir a distribuição em quatro partes iguais, quanto ao número de elementos em cada uma. Os critérios de classificação foram baseados na amplitude de CBE, e varia de zero a um ($0 \leq CBE \leq 1$). Assim, quanto mais próximo de 1,0 for o CBE, maior a classificação da estabilidade. A classificação como muito alta, significa que os genótipos e/ou ambientes, não sofrerão alterações significativas em seus postos, quando submetidos novamente a esses ambientes (genótipos) ou no conjunto desses genótipos (ambientes). Quanto à classificação como de alta estabilidade, espera-se não haver alterações significativas nos postos, porém, caso ocorram, que sejam de pequena magnitude, isto é, um e no máximo dois níveis. Embora, a estabilidade estatística dê garantia de que os postos não sofrerão alterações significativas, a qualidade dessa informação cai consideravelmente quando CBE aproxima-se de zero; assim, as classificações das estabilidades como média e baixa, requerem uma avaliação mais cuidadosa para recomendações desses genótipos e desses locais. Nesses casos, deve-se ponderar os riscos de uma indicação mais produtiva em relação a de maior constância nos resultados (Lavoranti, 2003).

As adaptabilidades, via *bootstrap* AMMI, foram avaliadas pelos produtos entre os escores *bootstrap* de genótipos (\mathbf{G}^*) e escores *bootstrap* de ambientes (\mathbf{H}^*), envolvendo apenas as duas primeiras colunas de \mathbf{G}^* e as duas primeiras linhas de \mathbf{H}^* . Os resultados das somas desses produtos indicaram as situações favoráveis para as respostas positivas e desfavoráveis para as respostas negativas (Lavoranti, 2003).

Vale ressaltar que as estabilidades de genótipos indicam boa adaptação aos ambientes de estudo. Assim, os sinais negativos para os produtos dos escores devem ser analisados com cautela, antes de serem descartados. O mesmo é válido para os genótipos instáveis que, apesar de sua adaptabilidade restrita, podem ser aproveitados vantajosamente na seleção e recomendação para um ambiente específico (Duarte e Vencovsky, 1999).

Todas as análises e gráficos foram implementadas pelo sistema estatístico SAS (SAS, 1997).

3 Resultados e discussão

Os escores AMMI têm distribuição de probabilidade desconhecida, fato que dificulta a obtenção de fórmulas via argumentos analíticos. Para medir a precisão desses escores, Lavoranti (2003) apresentou uma metodologia complementar para avaliar a precisão dos escores de genótipos e de ambientes, por meio de técnicas de amostragem *bootstrap*, permitindo uma melhoria na qualidade das inferências sobre as adaptabilidades e estabilidades fenotípicas estimadas pelo modelo AMMI.

Como os coeficientes dos escores de genótipos e ambientes são obtidos a partir da matriz de resíduos, procurou-se de acordo com Lavoranti (2003), associar a ferramenta estatística *bootstrap* às necessidades da análise AMMI. Em concordância com o modelo AMMI, que permite a estimação dos genótipos associados aos ambientes livres de ruídos, optou-se pela amostragem na matriz de resíduos, obtida dos valores estimados. Dessa forma, foram desenvolvidos procedimentos estatísticos que possibilitaram as análises gráficas e numéricas das adaptabilidades e estabilidades fenotípicas dos genótipos e dos ambientes envolvidos neste estudo.

Os resultados da análise da estabilidade via AMMI *bootstrap* de genótipos da população PCI (Figura 1) permitiram identificar como estáveis os genótipos que continham o ponto de coordenadas nulas (0,0) incluídas na região de confiança para o vetor *bootstrap* de médias. Todos os genótipos apresentaram estabilidade fenotípica, mas em graus variáveis sendo separados em quatro categorias distintas e citados a seguir em ordem decrescente de estabilidade, com base na magnitude dos coeficientes *bootstrap* apresentados na Tabela 1. No grupo de genótipos com estabilidade muito alta destacaram-se P3, P17, P16, P1, P6 e IAS-5; no grupo de alta estabilidade ficaram P5, IAC-100, P7, OCEPAR-4, P4, P8, P18 e P11; um terceiro grupo incluiu os seguintes genótipos com estabilidade média P13, P14, P20 e P2; finalmente, um quarto grupo de genótipos classificados como de estabilidade baixa foi formado por P10, P19, P9, P15 e P12.

Tabela 1 - Classificação pelo coeficiente *bootstrap* de estabilidade (CBE), de 20 linhagens (P1 a P20) da PCI (População com Controle Total de Insetos em F2) e três testemunhas comuns (T1: IAC-100, T2: Ocepar-4 e T3: IAS-5), para produtividade de grãos (kg/ha), em oito ambientes

Genótipos	N ⁽⁺⁾	CBE	Classificação	
P3	101	0,495	Estável	Muito alta
P17	112	0,440	Estável	Muito alta
P16	113	0,435	Estável	Muito alta
P1	115	0,425	Estável	Muito alta
P6	115	0,425	Estável	Muito alta
P5	117	0,415	Estável	Alta
P7	118	0,410	Estável	Alta
P11	119	0,405	Estável	Alta
P4	120	0,400	Estável	Alta
P8	120	0,400	Estável	Alta
P18	120	0,400	Estável	Alta
P13	121	0,395	Estável	Média
P14	121	0,395	Estável	Média
P20	122	0,390	Estável	Média
P2	123	0,385	Estável	Média
P10	124	0,380	Estável	Baixa
P19	126	0,370	Estável	Baixa
P9	127	0,365	Estável	Baixa
P15	128	0,360	Estável	Baixa
P12	143	0,285	Estável	Baixa
IAS-5	116	0,420	Estável	Muito alta
IAC-100	117	0,415	Estável	Alta
Ocepar-4	118	0,410	Estável	Alta

⁽⁺⁾: N é o número de pontos externos à região de predição para o vetor nulo.

Para a população formada a partir de plantas F_2 selecionadas para tolerância a insetos sugadores (PRIS), os resultados obtidos via construção da região de confiança *bootstrap* são mostrados na Figura 2. Todos os genótipos foram classificados como estáveis, sendo separados em quatro grupos com diferentes níveis de estabilidade, os quais estão a seguir relacionados em ordem decrescente de estabilidade (Tabela 2). O grupo com estabilidade muito alta envolveu os seguintes genótipos P10, P16, P9, P2, OCEPAR-4 e P5; entre as linhagens e testemunhas de alta presivibilidade ficaram P7, P8, P14, IAS-5, P6 e P11; genótipos medianamente estáveis compreenderam as linhagens P3, P13, P15, P17, P18, P1 e P12; já o grupo de baixa estabilidade incluiu os genótipos: P19, IAC-100, P4 e P20.

Tabela 2 - Classificação pelo coeficiente *bootstrap* de estabilidade (CBE), de 20 linhagens (P1 a P20) da PRIS (População com seleção para Resistência a Insetos Sugadores em F_2) e três testemunhas comuns (T1: IAC-100, T2: Ocepar-4 e T3: IAS-5) de soja, para produtividade de grãos (kg/ha) em oito ambientes

Genótipos	N ⁽⁺⁾	CBE	Classificação	
P10	112	0,440	Estável	Muito alta
P16	113	0,435	Estável	Muito alta
P9	114	0,430	Estável	Muito alta
P2	115	0,425	Estável	Muito alta
P5	116	0,420	Estável	Muito alta
P7	117	0,415	Estável	Alta
P8	118	0,410	Estável	Alta
P14	118	0,410	Estável	Alta
P6	119	0,405	Estável	Alta
P11	120	0,400	Estável	Alta
P3	121	0,395	Estável	Média
P13	121	0,395	Estável	Média
P15	121	0,395	Estável	Média
P17	121	0,395	Estável	Média
P18	122	0,390	Estável	Média
P1	123	0,385	Estável	Média
P12	123	0,385	Estável	Média
P19	124	0,380	Estável	Baixa
P4	129	0,355	Estável	Baixa
P20	130	0,350	Estável	Baixa
Ocepar-4	115	0,425	Estável	Muito alta
IAS-5	118	0,410	Estável	Alta
IAC-100	126	0,370	Estável	Baixa

(+): N é o número de pontos externos à região de predição para o vetor nulo.

Todos os genótipos da população PRIM apresentaram-se estáveis (Figura 3). Os genótipos foram classificados em quatro grupos distintos para o nível de estabilidade, estando relacionados a seguir em ordem decrescente de estabilidade (Tabela 3). O grupo com estabilidade muito alta incluiu os genótipos P16, P10, P8, P7, P5 e OCEPAR-4; os genótipos P2, P4, P11, P19, P9 e IAS-5 ficaram no grupo com alta previsibilidade; diagnosticou-se P18, P3, P6, P20, IAC-100 e P15 como genótipos medianamente estáveis; o grupo com baixa estabilidade envolveu as linhagens P1, P17, P14, P12 e P13.

Tabela 3 - Classificação pelo coeficiente *bootstrap* de estabilidade (CBE), de 20 linhagens (P1 a P20) da PRIM (População com seleção para Resistência a Insetos Mastigadores em F₂) e três testemunhas (T1: IAC-100, T2: Ocepar-4 e T3: IAS-5) de soja para produtividade de grãos (kg/ha) em oito ambientes

Genótipos	N ⁽⁺⁾	CBE	Classificação	
P16	108	0,46	Estável	Muito alta
P8	111	0,445	Estável	Muito alta
P10	110	0,45	Estável	Muito alta
P7	115	0,425	Estável	Muito alta
P5	116	0,42	Estável	Muito alta
P2	117	0,415	Estável	Alta
P4	119	0,405	Estável	Alta
P11	119	0,405	Estável	Alta
P19	119	0,405	Estável	Alta
P9	120	0,4	Estável	Alta
P18	121	0,395	Estável	Média
P3	122	0,39	Estável	Média
P6	122	0,39	Estável	Média
P20	122	0,39	Estável	Média
P15	123	0,385	Estável	Média
P17	124	0,38	Estável	Baixa
P1	124	0,38	Estável	Baixa
P14	127	0,365	Estável	Baixa
P12	128	0,36	Estável	Baixa
P13	128	0,36	Estável	Baixa
Ocepar-4	116	0,42	Estável	Muito alta
IAS-5	120	0,4	Estável	Alta
IAC-100	122	0,39	Estável	Média

(+): N é o número de pontos externos à região de predição para o vetor nulo.

Considerando os dois primeiros grupos de materiais genéticos com característica de muito alta e alta estabilidade fenotípica, a população PCI, apresentou maior número de

genótipos (60,87%) em comparação as populações PRIS e PRIM, ambas representando 52,17% de linhagens e testemunhas comuns constituindo esses dois grupos.

Os ambientes com alta estabilidade são aqueles que participaram em menor grau da interação de genótipos com ambientes, como foi evidenciado na Figura 4. Os oito ambientes foram classificados em quatro grupos distintos em ordem decrescente de estabilidade (Tabela 4). O ambiente E8 (Areão-CEI-01) apresentou menor dispersão de escores, sendo classificado como de estabilidade muito alta; em seqüência agrupam-se E7 (Areão-CII-01) e E4 (Areão-CEI-00) com alta estabilidade; com estabilidade baixa têm-se os ambientes E2 (Anhembí-CEI-00), E6 (Anhembí-CEI-01) e E5 (Anhembí-CII-01); um quarto grupo, representado pelos ambientes E1 (Anhembí-CII-00) e E3 (Areão-CII-00) foi classificado como instável; estes últimos ambientes não tiveram seus pontos de coordenadas nulos, redundando no aumento da variância da interação GxE. Os ambientes foram escolhidos por serem contrastantes entre si, o que possibilitou maior discriminação quanto à estabilidade, fato que os classificou como de muito alta estabilidade até instáveis. Muitas já estavam bastante homogeneizadas pelas sucessivas gerações de seleção, principalmente para produtividade, resistência a insetos e estabilidade fenotípica. Portanto, as linhagens não apresentam a classe instabilidade fenotípica.

Tabela 4 - Classificação pelo coeficiente *bootstrap* de estabilidade (CBE), de ambientes (E1 a E8), para produtividade de grãos (kg/ha) de populações (PCI, PRIS e PRIM) de soja

Ambientes	N ⁽⁺⁾	CBE	Classificação	
E8 – Areão-CEI-01	21	0,895	Estável	Muito alta
E7 – Areão-CII-01	68	0,660	Estável	Alta
E4 – Areão-CEI-00	128	0,360	Estável	Alta
E2 - Anhembí-CEI-00	131	0,345	Estável	Baixa
E6 - Anhembí-CEI-01	188	0,060	Estável	Baixa
E5 - Anhembí-CII-01	191	0,045	Estável	Baixa
E1 - Anhembí-CII-00	200	0	Instável	
E3 – Areão-CII-00	200	0	Instável	

(+): N é o número de pontos externos à região de predição para o vetor nulo.

Na análise de estabilidade pelo procedimento AMMI tradicional com representação gráfica em biplot, apenas apresentam falta de estabilidade fenotípica aqueles genótipos e/ou ambientes que tiveram as duas coordenadas (IPCA1 e IPCA2) distanciadas do zero ou muito distantes do IPCA1 (Lavoranti, 2003). Em relação a maior influência do IPCA1, de acordo com Johnson e Wichern (1998), a análise multivariada por componentes principais retem a maior porcentagem da variância explicada no primeiro eixo principal. Assim, a maior concentração do padrão de variação devido à interação, está concentrada no IPCA1 (Gauch e Zobel, 1988).

Houve uma tendência geral das regiões de confiança para estabilidade com níveis diferentes de inclusão do vetor nulo, mas não o suficiente para atestar possíveis instabilidades das linhagens. Visando melhorar a classificação das estabilidades, foram determinadas as regiões *bootstrap* de predição para o vetor nulo, e, a partir dos valores ordenados dos coeficientes *bootstrap* para a estabilidade, implementaram-se as

classificações das estabilidades fenotípicas de genótipos e ambientes via valores ordenados dos coeficientes *bootstrap*, pela análise numérica (Tabelas 1, 2, 3 e 4).

Apesar da versatilidade do modelo *bootstrap* AMMI, mais pesquisas devem ser conduzidas visando o pleno entendimento desse fenômeno, tão intrigante nos programas de melhoramento, que é a interação GxE. Um dos pontos que merecem estudos mais detalhados é a definição dos níveis de estabilidade, via região de predição para o vetor nulo, como, por exemplo, a análise gráfica da dispersão dos escores *bootstrap* (Lavoranti, 2003).

Um outro aspecto de maior relevância, representando um recurso adicional para melhor discernimento da metodologia aplicada, refere-se à avaliação dos gráficos da dispersão dos escores *bootstrap* de genótipos (Figuras 1, 2 e 3) e de ambientes (Figura 4). Aparentemente, neste processo, suspeita-se que o IPCA1 tem maior influência na classificação da estabilidade dos materiais, como pode ser evidenciado no gráfico das linhagens P12 (Figura 1) e P20 (Figura 2) e do ambiente E3 (Figura 4) identificados como genótipos de baixa estabilidade e ambiente instável, para os quais percebe que a dispersão da nuvem de pontos ocorreu ao longo do eixo IPCA2. Aquilantou-se, também, uma certa uniformidade para esses pontos de dispersão de escores entre os grupos de genótipos de muito alta, alta, média e baixa estabilidade, fato este que não se repetiu para os ambientes. Portanto, pode-se inferir com base nesse critério, que o descarte de genótipos considerados de baixa estabilidade via região de predição para os escores dos genótipos merece ser repensado. Cabe salientar que estas regiões de predição foram geradas com alto nível de confiança (99%) e que o caráter produtividade de grãos, por ser muito abalado por condições extrínsecas não controláveis, pode exercer influência a ser considerada quanto à escolha do nível de confiança fixado.

O padrão relativamente homogêneo de dispersão de escores entre os quatro grupos de genótipos estáveis tendeu a se repetir para as três populações, o que remete a proceder com cuidado sobre os critérios de eliminação de genótipos classificados aqui com baixa constância de comportamento mediante as variações ambientais.

Uma outra constatação que pode ser tirada situa-se no tamanho e forma da região de confiança, estando esta correlacionada com a nuvem de dispersão dos pontos dos escores, delineada predominantemente no tipo elíptico, com diâmetro maior na horizontal quando prevalece concentração dos pontos dos escores no sentido do IPCA1; todavia, quando esta nuvem de pontos está mais concentrada ao longo do IPCA2, o desenho elíptico tem maior diâmetro vertical; se houver uma distribuição de pontos que contemple simultaneamente os dois eixos de forma semelhante, será plotado no gráfico um círculo. Essa associação é definida também pelos valores de desvio padrão para cada eixo, sendo maior o diâmetro quanto maior for o desvio padrão do eixo respectivo, isto é, o diâmetro da região de confiança *bootstrap* acompanha a tendência da magnitude do desvio padrão. Quando a elipse é do tipo horizontal entende-se que o IPCA1 dominou, isto é, maior porção da variação foi retida neste eixo, como o esperado. Quanto menor o tamanho dessas regiões e mais próximos estiverem do ponto de coordenadas (0,0) que inclui o vetor nulo, espera-se manifestação mais alta de estabilidade dos materiais avaliados. Esse deslocamento da região de confiança ditado pelo padrão da nuvem de pontos é facilmente identificado na Figura 4. Nessa avaliação, observou-se que os genótipos apresentaram menor oscilação do que os ambientes, confirmando as magnitudes dos testes F para esse efeito.

A soma de quadrados da interação GxE ($SQ_{G \times E}$) é importante para explicar o padrão adjacente à interação GxE. Assim, quando se cometeu o erro tipo II ao aceitar um modelo

AMMI com menor número de eixos, cujo modelo correto, seja de fato mais parametrizado, foi excluído parte da $SQ_{G \times E}$, que contém variações devidas aos padrões de interação $G \times E$, que são de interesse agrônomo. Dessa forma, as interpretações de estabilidade ficam comprometidas. Para evitar os riscos desse tipo de erro, recomenda-se utilizar outras técnicas mais acuradas para a definição do modelo, como as apresentadas por Dias e Krzanowski (2003), combinando-se essas técnicas (seleção de modelos “exatos” e *bootstrap* AMMI).

O procedimento AMMI *bootstrap* teve caráter de refinamento para análises de estabilidade fenotípica, mostrando grande potencial complementar à análise AMMI, apresentando recursos adicionais para ajudar na seleção mais criteriosa de genótipos e ambientes para estabilidade fenotípica.

Conclusões

As seguintes conclusões foram obtidas, com base na produtividade de grãos:

- o método AMMI com reamostragens (*bootstrap*) mostrou resultados mais detalhados do que os obtidos apenas com o método AMMI;
- para as três populações, foram formados quatro grupos de genótipos com diferentes níveis de estabilidade, tendo a soma dos dois grupos de estabilidade mais elevada incluído pelo menos 50% das linhagens de cada população e duas das três testemunhas;
- as linhagens foram classificadas como estáveis, com graus variáveis de estabilidade fenotípica;
- os ambientes apresentaram as classes extremas de estabilidade fenotípica: máxima estabilidade e ambientes instáveis.

MAIA, M. C. C.; VELLO, N. A.; LAVORANTI, O. J.; DIAS, C. T. dos S.; VENCOVSKY, R.; ROCHA, M. de M.; NUNES, G. H. de S. AMMI-*Bootstrap* in the study of genotype x environment (GxE) interaction in soybean. *Rev. Mat. Est.*, São Paulo, v.24, n.3, p.7-24, 2006.

- ABSTRACT:** *This study aimed to investigate the genotype x environment (GxE) interaction and phenotypic stability by applying AMMI (Additive Main Effects and Multiplicative Interaction Model) and bootstrap methods for evaluating seed yield of soybean lines. The material included three populations derived from a 4x4 partial diallel having four insect resistant parents and other four parents with high seed yield and earliness. From F_3 to F_{11} , selection was made from progenies for seed yield and photoperiod tolerance. The 24 experiments were carried out in randomized block design with two replications subdivided in sets with common controls. Lines of the three populations were evaluated in eight environments (E1 to E8): two locations (Anhembí, Areão), two years (1999/00, 2000/01), and two management systems (CII: intensive control of insects; CEI: ecological control of insects). The following conclusions could be drawn: all lines were classified as stable but with different stability levels. The two groups with higher stabilities included jointly at least 50% of the lines and two checks. The (AMMI + bootstrap) method detected E8 (Areão, 2000/01, CEI) as the most stable environment; on the other hand, E1 (Anhembí, 1999/00, CII) and E3 (Areão, 1999/00, CII) were classified as unstable environments.*

The AMMI method with bootstrapping led to more detailed results than the AMMI method without resampling.

- **KEYWORDS:** *AMMI-bootstrap; genotype x environment interaction; glycine max (L.)*

REFERÊNCIAS

BRASIL, E. M. *Comparação de métodos no estudo da interação de genótipos com ambientes em milho (Zea mays L.)*. Goiânia, 1990. 181f. Dissertação (Mestrado em Agronomia) – Universidade Federal de Goiânia, Goiânia, 1990.

DIAS, C. T. S.; KRZANOWSKI, W. J. Model selection and cross-validation in additive main effect and multiplicative interaction (AMMI) models. *Crop Sci.*, Madison, v.43, n.3, p.865-873, 2003.

DUARTE, J. B.; VENCOVSKY, R. *Interação genótipos x ambientes: uma introdução à análise AMMI*. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 60p. (Série Monografias, 9).

DUARTE, J. B.; ZIMMERMANN, M. J. O. Correlation among yield stability parameters in common bean. *Crop Sci.*, Madison, v.35, n.3, p.905-912, 1995.

GAUCH, H. G.; ZOBEL, R. W. Predictive and postdictive success of statistical analysis of yield trials. *Theor. Appl. Genet.*, Berlin, v.76, n.1, p.1-10, 1988.

GAUCH, H. G.; ZOBEL, R. W. AMMI analysis of yield trials. In: KANG, M. S.; GAUCH, H. G. (Ed.) *Genotype-by-environment interaction*. New York: CRC Press, 1996. p.85-122.

JOHNSON, R. A.; WICHERN, D. W. *Applied multivariate statistical analysis*. Madison: Prentice Hall International, 1998, 816p.

LAVORANTI, O. J. *Estabilidade e adaptabilidade fenotípica através da reamostragem "Bootstrap" no modelo AMMI*. 2003. 166f. Tese (Doutorado em Estatística e Experimentação Agrônômica) – Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2003.

MANDEL, J. A new analysis of variance model for non-additive data. *Technometrics*, Washington, v.13, n.1, p.1-18, 1971.

STATISTICAL ANALYSIS SYSTEM INSTITUTE. *SAS/STAT software: changes and enhancements through release 6.12*. (software). Cary: SAS Institute, 1997. 1116p.

ZOBEL, R. W. et al. Statistical analysis of a yield trial. *Agron. J.*, Madison, v.80, n.3, p.388-393, 1988.

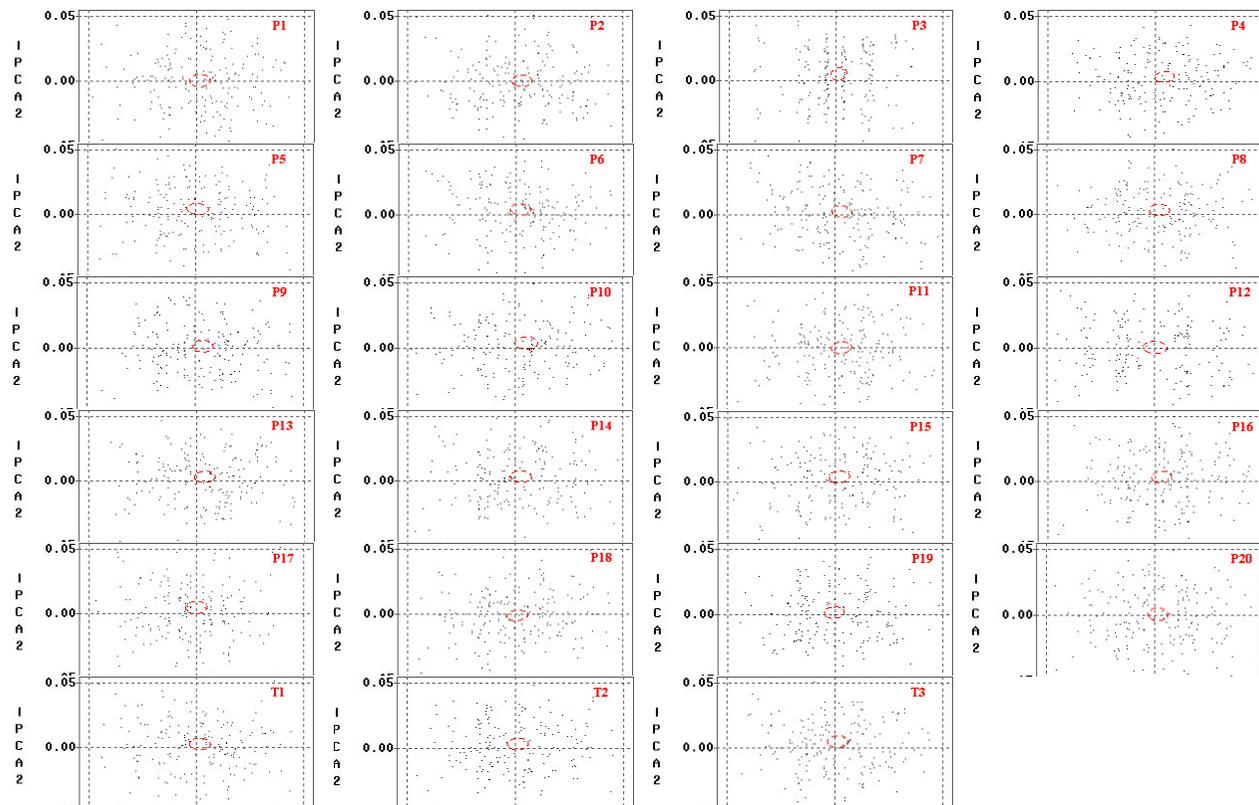


Figura 1- Gráficos de dispersão e regiões de confiança de 99% para os escores “bootstrap” de 20 linhagens (P1 a P20) da PCI (população com **Controle Total de Insetos** em F_2) e três testemunhas (T1: IAC-100, T2: Ocepar-4 e T3: IAS-5) de soja, para produtividade de grãos (kg/ha) em oito ambientes.

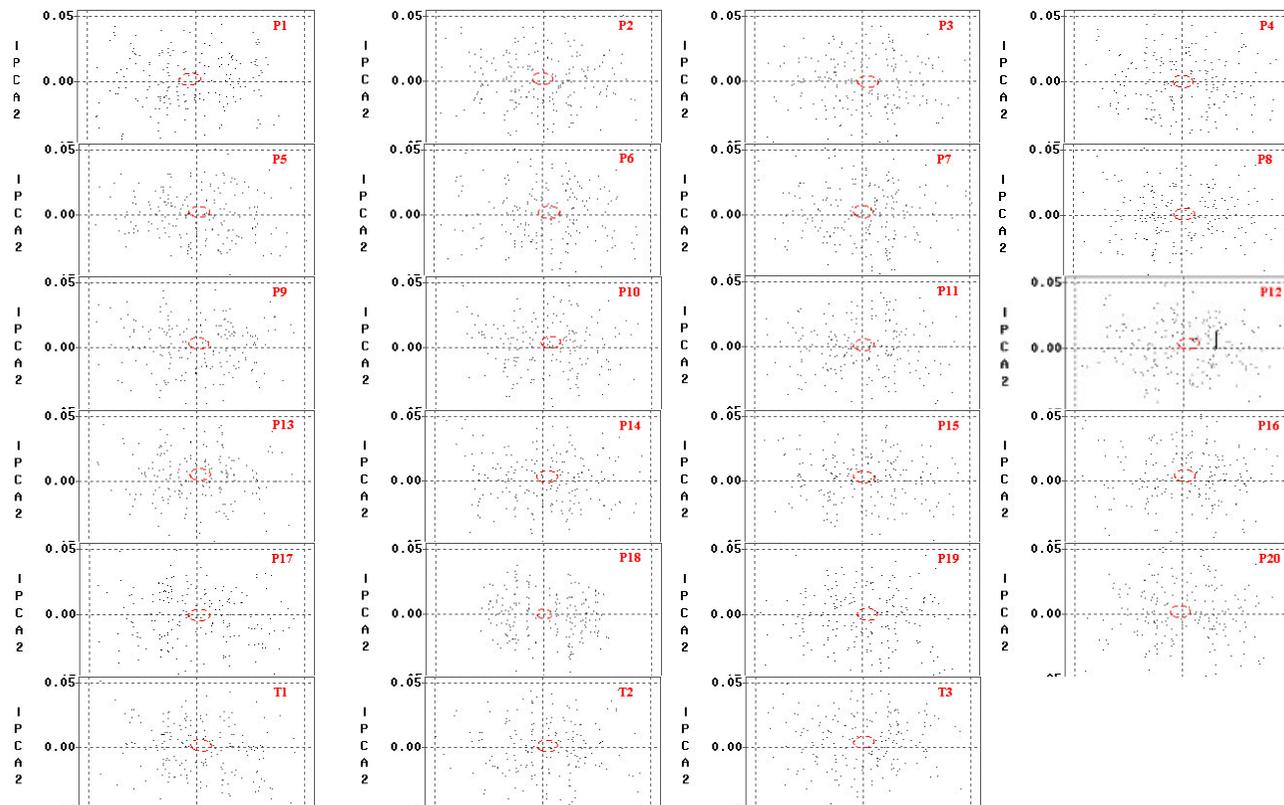


Figura 2- Gráficos de dispersão e regiões de confiança de 99% para os escores “bootstrap” de 20 linhagens (P1 a P20) da PRIS (População com seleção para Resistência a Insetos Sugadores em F_2) e três testemunhas (T1: IAC-100, T2: Ocepar-4 e T3: IAS-5) de soja, para produtividade de grãos, (kg/ha) em oito ambientes.

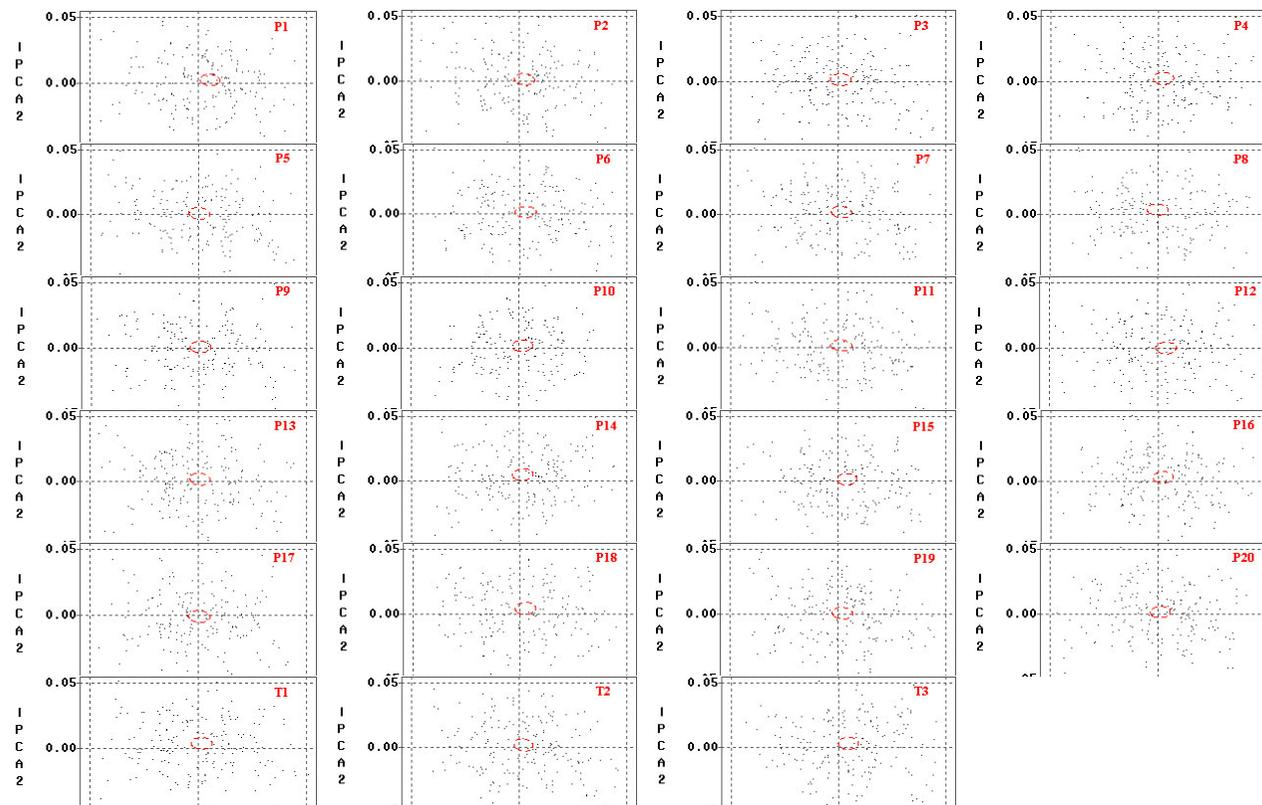


Figura 3- Gráficos de dispersão e regiões de confiança de 99% para os escores “bootstrap” de 20 linhagens (P1 a P20) da PRIM (População com seleção para Resistência a Insetos Mastigadores em F2) e três testemunhas (T1: IAC-100, T2: Ocepar-4 e T3: IAS-5) de soja para produtividade de grãos (kg/ha) em oito ambientes.

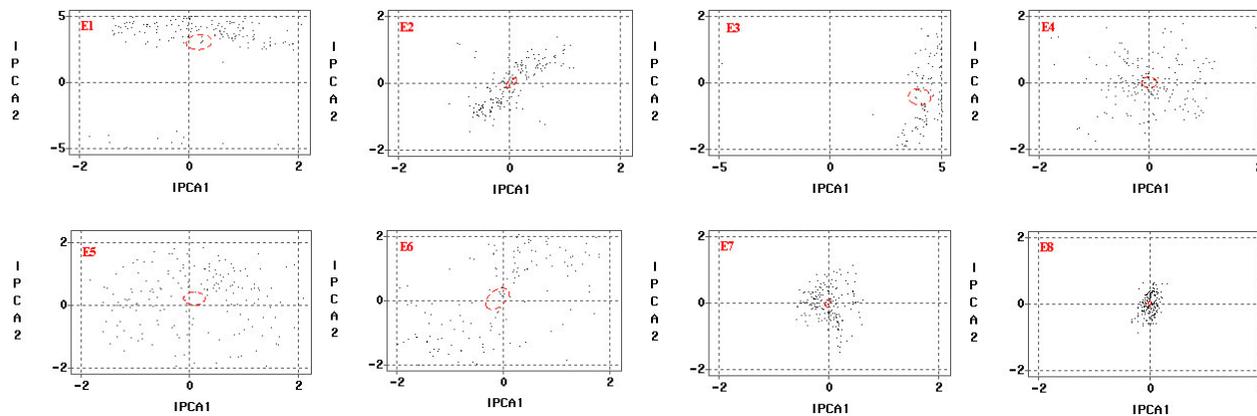


Figura 4 - Gráficos de dispersão e regiões de confiança de 99% para os escores “bootstrap” de ambientes (E1 a E8), para produtividade de grãos (kg/ha), de três populações (PCI, PRIS e PRIM) de soja

Recebido em 26.04.2006.
Aprovado após revisão em 11.08.2006.